R E P O R T  R E S U M E S

ED 012 494 CG 000 659

USEFULNESS OF PEER RATINGS OF PERSONALITY IN EDUCATIONAL RESEARCH.

BY- SMITH, GENE M.

EDRS PRICE MF-$0.25 HC-$0.88 22P.

DESCRIPTORS- *EDUCATIONAL RESEARCH, *PERSONALITY STUDIES, *ACADEMIC ACHIEVEMENT, COLLEGE STUDENTS, HIGH SCHOOL STUDENTS, NURSES, *PEER RELATIONSHIP, PREDICTIVE MEASUREMENT, PERSONALITY TESTS, PEER RATING TECHNIQUES

ALTHOUGH PEER RATINGS WERE NOT REGARDED AS VALID INDICATORS IN TEST STUDIES, THIS PROJECT DEMONSTRATES THAT PEER RATINGS OF PERSONALITY CAN BE HELPFUL IN CLARIFYING THE RELATIONSHIP BETWEEN PERSONALITY AND ACADEMIC SUCCESS. TEST CONDITIONS WERE DESIGNED TO AVOID METHODOLOGICAL PROBLEMS. PERSONALITY VARIABLES WERE CAREFULLY ANALYZED, AND A FORCED-CHOICE PROCEDURE WAS ADOPTED. TEST SUBJECTS WERE CAPABLE OF EFFECTIVELY RATING ONE ANOTHER BECAUSE THEY INTERACTED FREQUENTLY. PEER RATINGS WERE COLLECTED BEFORE THE FIRST MIDTERM EXAMINATIONS IN ORDER TO PREVENT AN ACADEMIC PERFORMANCE BIAS. THE RESULTS OF RELIABILITY AND FACTOR ANALYTIC STUDIES PERFORMED ON PEER-RATING DATA DEMONSTRATED A POSTIVE RELATIONSHIP BETWEEN PERSONALITY AND ACADEMIC SUCCESS. THE STUDY SHOWS THAT PERSEVERANCE, CONSCIENTIOUSNESS, INQUISITIVENESS, RESPONSIBILITY, SELF-RELIANCE, AND ORDERLINESS ARE RELATED TO ACADEMIC SUCCESS IN THE POPULATION STUDIED. ALL THESE TRAITS ARE NONINTELLECTIVE CORRELATES OF ACADEMIC SUCCESS BELONGING TO THE SAME FACTOR, "STRENGTH OF CHARACTER." THIS FACTOR WAS FOUND TO HAVE THE HIGHEST PREDICTIVE VALIDITY OF ALL PEER VARIABLES. (AF)

# Usefulness of Peer Ratings of Personality in Educational Research

Gene M. Smith, Harvard Medical School and Boston University

For more than 30 years researchers have sought to clarify the relationship between nonintellective factors and academic success. The problem is of basic theoretical and methodological significance and has national practical importance, but progress toward its solution has been slow (Fishman and Pasanella, 1960; Gaier and White, 1965; Garrett, 1949; Harris, 1940; Michael, 1965; Stagner, 1933; Travers, 1949). Most psychologists and educators would agree that poor motivation, faulty attitudes, and other nonintellective problems contribute importantly to academic failure. Yet evidence to support this agreement is elusive and based largely on consideration of individual cases.

If the relationship between nonintellective factors and academic success is to be elucidated further, it appears that new approaches must be tried. The results presented here, and in a concurrently published report (Smith, 1967), indicate that peer ratings of personality (an approach to personality assessment not often used by psychologists) can be remarkably helpful in clarifying the relations between personality and academic success. Another concurrently published report (Smith, 1966), dealing with an entirely different problem, gives further evidence of the usefulness of peer rating data for analysis and solution of theoretical and empirical problems.

Although several studies (e.g., Astington, 1960; Carroll, 1952; Doll, 1963; Flyer, 1963; Flyer and Bigbee, 1954; Kleiger et al., 1962; Tupes, 1957) have shown peer ratings of personality to have good reliability and predictive validity, this method of studying personality has never been widely appreciated--perhaps, in part, because of biases which can reduce the validity of rating data (Guilford, 1954, Secord, 1958, and Guilford et al., 1962) and, in part, because of the difficulty of achieving test conditions required for the valid use of the peer rating method.



CG 000 659

The purposes of this paper are (a) to discuss test conditions needed to avoid methodological problems frequently encountered in use of rating data, (b) to report results of reliability and factor analytic studies performed on peer rating data obtained in samples from college, nursing school, and high school and (c) to present results concerning relations between personality and academic performance in a sample of 348 college students.

## METHOD

The Peer Rating Technique Used in this Study. The rater examines each of 42 bipolar personality traits and selects the five members of his peer group most like the left hand pole and the five most like its opposite on the right. (See example in fig. 1.) Selections on the left are considered positive nominations and those on the right are considered negative. The positive and negative nominations a ratee receives are scored +1 and -1, respectively. Failure to be nominated is scored zero. Thus, in a peer group of 25, a ratee's score on any trait can range from +24 to -24.

Three features of this method should be emphasized: (1) For each trait a rater must nominate five peers for the left pole and five for the right; if he does not comply at least 39 times out of 42, his ratings are discarded. (2) Since a ratee is rated on each of 42 traits by each of 20-30 peers, he receives about 1,000 separate ratings from peers with whom he has a high frequency, intensity, and duration of social interaction. (3) The 42 traits on which a peer is rated are the product of extensive work: Allport and Odbert (1936) compiled all words in the English dictionary (1,800) which might be used to describe personality; Cattell (1957), by a series of logical and empirical distillations, reduced these traits to 42 rating items.

Guilford (1954) emphasizes the dangers inherent in rating techniques: Error can arise from (1) individual rater response tendencies (e.g., errors of leniency, neutrality, overseverity, contrast, and similarity), (2) the unique

relationship between the rater and the ratee (e.g., errors of halo and irrelevance), (3) inadequate knowledge of the ratee by the rater, and (4) faulty construction of the items to be rated.

Appropriate planning can reduce the effects of difficulties such as those just mentioned. The errors of leniency, neutrality, overseverity, contrast, and similarity, which arise when the rater can assign the various positions on a rating scale without restriction, reflect the rater's response preference. Use of a forced choice procedure allows the rater no response preference.* The rating instructions, used to collect the data reported here, end by saying:

> "On each card you must fill in five numbers on the left and five on the right. This will be difficult in some cases. For instance, in the case of "adaptable" vs. "rigid", you might feel that most members of your section are adaptable and only two or three are rigid. You still have to fill in five numbers under "rigid", and you would do that by selecting the five you think are least adaptable."

Invalid ratings due to errors of halo and irrelevance are due to the rater's subjective bias. Unless many raters have the same bias, use of many raters should tend to cause these individual biases to cancel. (The present study of college students employed 20-30 raters per group, a number which exceeds that usually found in studies employing rating scales.)

Rating data have been criticized as undependable because of insufficient knowledge of ratees by raters. This criticism usually arises when one or a few superiors rate subordinates whom they observe playing one role in one environmental setting. In the present study, the subjects were peers; there were many (usually 15-30) raters of each peer; and the subjects interacted frequently, over long periods, and under diverse environmental circumstances.

Regarding the question of faulty construction of items to be rated, three points should be made. (a) As already stated, the derivation of the 42 items

* It is important to note that forced choice among ratees is conceptually and methodologically different from forced choice between traits presumed to be equated for social desirability. There is evidence that with the latter method, biases of the type under discussion can occur (Lovell & Haner, 1955; Travers, 1951).

is based on considerable work by Allport and Odbert (1936) and by Cattell (1957). (b) The reliability and predictive validity of these items is high. (c) Most of these items appear to meet the exacting standards specified by Guilford (1954) and originally published by Champrey (1941): clarity, relevance, precision, variety, objectivity, and uniqueness. (See listing of 42 items; Cattell, 1957, pp. 813-817)

Length of Acquaintance in the College Sample. In studies where peer ratings of personality are collected after students in a peer group have interacted for a long period of time, it is possible for correlations between academic performance and peer ratings of personality to be biased by a tendency for academic performance to influence peer ratings. To reduce the effect of this potential bias in the study of the college sample reported here, peer ratings were collected before the 348 college students received results of their first midterm examinations. (At that point they had been together only nine weeks.)

Other Characteristics of the College Sample. The 348 college students entered the College of Basic Studies of Boston University (CBS) in 1964*. CBS is a two-year lower division program which employs team teaching and a core curriculum. An entering class of approximately 550 freshmen is divided into 20 sections of 25-30 students each, four sections of which are assigned to a team of five instructors who represent the five divisions which make up the core curriculum of the College--Humanities, Science, Social Science, Rhetoric, and Psychology and Guidance. CBS admits applicants who (because of poor high school records, shortages in prerequisites, low academic aptitude scores, or some combination) are denied admission into four-year programs in Boston

* 235 of the 583 students in this class were excluded from analyses reported in tables 2-4 because they lacked scores on the criterion (year one grade point average) or on one or more of the variables numbered 48-77 in table 2. All 583 had scores on the peer rating variables and thus were used in the factor analysis of data of college students shown in table 1.

University. The CBS program is described in further detail by Anthony et al., 1956, LaFauci, 1965, and LaFauci & Richter, 1965.

The joint use of team system and core curriculum causes CBS students in sections of 25-30 students to have a frequency, intensity, and length of social interaction rarely found in other collegiate programs. Although these features make CBS a very suitable population in which to use the peer method, many other academic settings also provide suitable samples; e.g., high schools where sectioning by curriculum, and by ability within curriculum, results in students in groups of 20-30 having many classes together.

## RESULTS

Reliability of Peer Ratings of Personality. Data collected at CBS, one high school, and two nursing schools, have permitted the study of split-half reliability of peer ratings on Cattell's 42 bipolar behavior traits. Figure 2 shows, for each of 71 samples, the median of the 42 split-half reliability coefficients (corrected with the Spearman-Brown formula) plotted against the number of raters in that sample. (Variation among samples regarding number of raters was due partly to variation in size of peer groups and partly to variation among groups regarding absenteeism on the test day and regarding compliance with rating instructions.) Note that reliability varies as a function of number of raters, increasing as raters increase from six to about twenty. After twenty, increased reliability is slight. The median of the median reliability coefficients obtained with samples of 15 or more raters was 0.83; see figure 2.

Factor Analytic Studies. Factor analysis (principal components rotated to an oblique solution using the biquartimin criterion, Harmann, 1960) of scores on the 42 traits typically yields five factors which have stable structure from sample to sample within the CBS population and which agree well with results obtained with other populations we have studied (high school and nursing students) and with populations studied by others. The five factors we find are similar to

those reported by Tupes and Christal (1961) and by Norman (1963). The terms we use to designate the five factors are "agreeableness", "extraversion", "strength of character", "emotionality", and "refinement". Table 1 shows the loadings obtained on each of these five factors in each of three large samples: 583 college students, 521 nursing students, and 324 high school students. (These three samples contain 20, 22, and 14 peer groups, respectively; most peer groups consisted of 15-30 members.) The peer variables are identified by number in table 1 and by number and name in table 2. (See Cattell, 1957, p.813-817 for full definitions of the 42 variables.) The percents of trace listed at the bottom of table 1 show that the relative importance of the five factors (in terms of variance accounted for) is the same in all three samples. The factor analytic structure of the 42 peer variables, shown in table 1, is psychologically meaningful and very similar for the three samples.

Prediction of Academic Success. In both univariate and multivariate studies of data obtained from the 348 college students, the predictive validity of the peer variables belonging to the factor we call "strength of character" was found to be superior to that of variables belonging to the other four peer factors and superior to that of the 30 non-peer variables (i.e., 13 measures of academic aptitude, two measures of high school performance, and 15 scores on the EPPS scales)*.

In univariate studies, the peer variables which predicted year 1 grade point average best were: "quitting" (-.47), "inquisitive" (+.35), "conscientious"

* The 13 measures of academic aptitude were: the verbal and mathematical subtests of the College Entrance Examination Board Scholastic Aptitude Test, five subtests of the Differential Aptitudes Test (verbal reasoning, numerical reasoning, abstract reasoning, mechanical reasoning, and space relations), four subtests of the Cooperative English Test (vocabulary, reading speed, reading level, and mechanics of expression), the writing subtest of the Sequential Tests of Educational Progress Series, and the Watson-Glaser Critical Thinking Appraisal (See Buros, 1959). The two high school measures were: high school rank corrected for class size [i.e., 1.00 - (rank/size)], and number of certified high school units.

(+.32), "prone to daydream" (-.32), "responsible" (+.29), "insistently orderly" (+.28), "self-reliant" (+.26), "languid" (-.26), "socially mature" (+.25), and "resourceful" (+.24). Nine of these 10 variables belong to the factor we call "strength of character". A composite score for "strength of character" (based on all variables belonging to this factor) correlated +.43 with the criterion. (See table 2.)

Among the 13 academic aptitude measures, the STEP-Writing, Cooperative English Reading Level, Cooperative English Vocabulary, and SAT-Verbal gave the highest univariate correlations with year 1 grade point average (+.25, +.23, +.22, and +.22, respectively). High school rank and number of certified units yielded correlation coefficients with the criterion of +.19 and +.18 respectively. (See table 2.)

In multiple regression analysis (Rao, 1952) of the data of the 348 college students, the contribution to predictive accuracy made by the peer data exceeded that of all other sources of data combined. Prediction of year 1 grade point average from a battery consisting of 72 variables (42 peer variables, 13 academic aptitude variables, two high school performance variables, and 15 EPPS variables) was performed with a computer program in which variables were admitted to the prediction battery in a "step-wise" manner in the order of decreasing contribution to predictive accuracy. In the procedure used, the first variable to enter the prediction equation was the one (out of 72) which had the highest correlation with the criterion. The next to enter was the one having the highest correlation with the criterion after the effect of the first predictor variable was partialled out. At the next and at each succeeding step the entering variable was the one having the highest correlation with the criterion when the influence of the multivariate prediction battery obtained at the preceding step of the analysis was partialled out.

In the results reported in tables 3 and 4, the analyses were stopped after 10 variables* had entered the prediction battery. At the tenth step of the analysis reported in table 3 the multiple correlation coefficient was .64 and the contribution of each of the 10 variables was statistically significant. Six of these 10 were peer variables, two were aptitude variables, and two were variables dealing with high school performance. The total contribution to $R^2$ made by the four types of variables in this analysis were: peer 68%, aptitude 19%, high school 13%, and Edwards zero%. The peer variable called "quitting" (see figure 1 for complete definition) contributed more to predictive accuracy than any of the other 71 predictor variables in this analysis. It accounted for 55% of $R^2$. As shown in table 1, "quitting", which is peer variable #5, belongs to the factor we call "strength of character".

A parallel analysis was conducted by substituting five composite** peer variables for the 42 original peer variables. Under those conditions the multiple correlation coefficient reached .60 at the tenth step of the analysis. The total contribution to $R^2$ was: peer 59%, aptitude 18%, high school 16%, and Edwards 6%. The composite peer variable we call "strength of character" contributed more to predictive accuracy (It accounted for 52% of $R^2$) than any of the other 24 predictor variables (See table 4.).

## COMMENTS

The sample of students for whom results are reported in tables 2, 3, and 4 is one of several samples studied at CBS and elsewhere with the peer rating

* In such analyses we usually find that predictive accuracy is not increased appreciably by allowing more than six or eight variables to enter the equation. Hence, to conserve computer time, the analyses were stopped after the tenth step of the equation.

** A composite is an average of scores on variables comprising a factor. (See table 1 for composition of composites.) The composite score for "emotionality", for example, is the average of scores on variables 2,19,24, and 27 in table 1, where 2,19, and 27 are given coefficients of +1 and 24 is given a coefficient of -1.

method of assessing personality. Results obtained with this particular sample are of special interest because in it the peer data were collected early in the acquaintance of the peers (just prior to report of midterm grades for the first semester of the freshman year and hence six months before data comprising the criterion--year 1 grade point average--were generated) in an attempt to avoid bias due to rater knowledge of academic performance of ratees. Both reliability and predictive validity of the "short acquaintance" peer data were somewhat lower than when CBS peer data were collected after longer acquaintance (e.g., the median reliability of the short acquaintance peer data was .69, whereas that for CBS samples studied after one year of acquaintance is usually between .80 and .85) but the particular peer variables which had the highest predictive validity in the present analyses (namely, those belonging to what we call "strength of character"; see tables 1 and 2) are those which other unpublished analyses have shown to be most highly correlated with academic success in CBS samples studied after longer acquaintance.

The CBS population is unique in the respects mentioned in the last two paragraphs of METHOD. In addition, the univariate correlations between the criterion (year 1 grade point average) and typically useful predictors such as high school performance and academic aptitude scores were somewhat lower than is usually found in the literature. Nevertheless, there are reasons for believing the relationships found in this study are not exclusively indigenous to CBS. In concurrently published studies of students from nursing school and high school (Smith, 1967), the results obtained with peer data are similar to those obtained in the CBS study. Further work is needed to establish the generality of these findings, but it should be mentioned that peer rating measures of perseverance, conscientiousness, inquisitiveness, responsibility, self-reliance, and orderliness found to be related to academic success in the CBS population have also

been found to be related to academic success in analyses of data from other populations we have studied. It is of particular interest that all of these traits belong to the same factor, called here "strength of character".

## SUMMARY AND CONCLUSIONS

The relationship between personality and academic success, though recognized as important, is not well understood. This paper presents results indicating that peer ratings of personality (an approach to personality assessment not often used by psychologists) can be remarkably helpful in clarifying this elusive relationship. Over 1,000 students from college, high school, and nursing school contributed data reported here. Results are presented concerning the reliability and factor analytic structure of peer data obtained in all three populations studied, and concerning the predictive validity of peer and non-peer data obtained in one sample of 348 college students. In addition, the paper discusses ways of avoiding methodological problems frequently encountered in use of rating data.

Three conclusions are drawn: (1) Peer ratings of personality, properly elicited and evaluated, can provide information of high reliability and predictive validity. (2) The factor analytic structure of the 42 personality variables studied with the peer rating technique is highly stable from sample to sample within and across populations. (3) The peer variables belonging to the factor called here "strength of character" are important nonintellective correlates of academic success.

## REFERENCES


Allport, G.W., & Odbert, H.S. Trait-names: A psycholexical study. Psychological Monograph, 1936, 47, 171.

Anthony, V.A., Livesey, C., Richter, P.E., & Russell, C.H. The team approach to general education. Junior College Journal, 1956, 26, Part I, 319-27, Part II, 405-10.

Astington, E. Personality assessment and academic performance in a boy's grammar school. British Journal of Educational Psychology, 1960, 30, 225-36.

Buros, O.K. (Ed.) The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1959.

Carroll, J.B. Ratings on traits measured by a factored personality inventory. Journal of Abnormal and Social Psychology, 1952, 47, 626-632.

Cattell, R.B. Personality and motivation structure and measurement. Yonkers, New York: World Book Co., 1957.

Champrey, H. The measurement of parent behavior. Child Development, 1941, 12, 131-166.

Doll, R.E. Officer peer ratings as a predictor of failure to complete flight training. Aerospace Medicine, 1963, 34, 130-131.

Edwards, A.L. Personal Preference Schedule. New York: The Psychological Corporation, 1954.

Fishman, J.A., & Pasanella, A.K. College admission-selection studies. Review of Educational Research, 1960, 30, 298-310.

Flyer, E.S. Prediction of unsuitability among first-term airmen from aptitude indexes, high school reference data, and basic training evaluations. (PRL-TDR-63-17), Lackland Air Force Base, Texas: Air Force Personnel Research Laboratory, June 1963.

Flyer, E.S., & Bigbee, L.R. The light plane as a pre-primary selection and training device: III, Analysis of selection data. Lackland Air Force Base, Texas: Air Force Personnel Research Laboratory, December, 1954.

Gaier, E.L., & White, W.F. Trends in the measurement of personality. Review of Educational Research, 1965, 35(1), 63-81.

Garrett, H.F. A review and interpretation of investigations of factors related to scholastic success in colleges of arts and science and teachers colleges. Journal of Experimental Education, 1949, 18, 91-138.

Guilford, J.P. Psychometric methods. New York: McGraw-Hill, 1954.

Guilford, J.P., Christensen, P.R., Taaffe, G., & Wilson, R.C. Ratings should be scrutinized. Educational and Psychological Measurement, 1962, 22, 439-447.

Harmann, H.H. Modern factor analysis. Chicago: University of Chicago Press, 1960.

Harris, D. Factors affecting college grades: A review of the literature, 1930-1937. Psychological Bulletin, 1940, 37, 125-166.

Klieger, W.A., deJung, J.E., Dubuisson, A.U. Peer ratings as predictors of disciplinary problems. Washington, D.C.: U.S. Army Personnel Research Office, 1962.

LaFauci, H.M. Team teaching in a general education program. Journal of General Education, 1965, 17(2), 149-159.

LaFauci, H.M., & Richter, P.E. Unity and variety in a general education program. Journal of Higher Education, 1965, 36(7), 379-384.

Lovell, G.D., & Haner, C.F. Forced-choice applied to college faculty rating. Educational and Psychological Measurement, 1955, 15, 291-304.

Michael, W.B. A short evaluation of the research reviewed in educational and psychological testing. Review of Educational Research, 1965, 35(1), 92-99.

Norman, W.T. Toward an adequate taxonomy of personality attributes: replicated factor structure in peer nomination personality ratings. Journal of Abnormal and Social Psychology, 1963, 66, 574-583.

Rao, C.R. Advanced statistical methods in biometric research. New York: Wiley, 1952.

Secord, P.F. Facial features and inference processes in interpersonal perception. In R. Tagiuri and T. Petrullo (Eds.), Person perception and interpersonal behavior. Stanford, California: Stanford University Press, 1958. Pp. 300-315.

Smith, G.M. Personality correlates of cigarette smoking in students of college age. Annals of the New York Academy of Sciences, in press.

Smith, G.M. Usefulness of peer ratings for prediction of academic success of students in nursing school and high school. (Submitted to Nursing Research.).

Stagner, R. The relation of personality to academic aptitude and achievement. Journal of Educational Research, 1933, 26, 648-655.

Travers, R.M.W. Significant research on the prediction of academic success. In W.T. Donahue and associates (Eds.), The measurement of student adjustment and achievement. Ann Arbor: University of Michigan Press, 1949.

Travers, R.M.W. A critical review of the validity and rationale of the forced-choice technique. Psychological Bulletin, 1951, 48, 62-70.

Tupes, E.C. Relationships between behavior trait ratings by peers and later officer performance of USAF officer candidate school graduates. Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, October, 1957.

Tupes, E.C., & Christal, R.E. Recurrent personality factors based on trait ratings. USAF WADC Tech. Rept. No. 61-97, Dayton Ohio: Wright Air Development Center, 1961.

Table 1

Loadings of 42 Peer Rating Variables on Each of Five Factors in Studies of 583 College, 521 Nursing and 324 High School Students

| Factors: | Agreeableness | | | Extraversion | | | Strength of Character | | | Emotionality | | | Refinement | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Samples: | Col. | Nur. | H.S. | Col. | Nur. | H.S. | Col. | Nur. | H.S. | Col. | Nur. | H.S. | Col. | Nur. | H.S. |
| 1. | +.70 | +.73 | +.63 | +.10 | +.29 | +.08 | -.04 | +.02 | +.08 | -.23 | -.34 | -.10 | -.01 | +.06 | +.06 |
| 4. | -.74 | -.82 | -.66 | +.06 | +.01 | -.03 | +.04 | -.02 | -.07 | +.23 | +.20 | +.18 | +.16 | +.14 | +.01 |
| 6. | +.65 | +.82 | +.63 | +.05 | +.01 | +.08 | +.04 | -.14 | +.16 | +.34 | +.21 | +.28 | +.06 | +.14 | +.05 |
| 7. | +.69 | +.79 | +.67 | -.25 | -.19 | -.08 | -.01 | -.14 | +.06 | +.07 | +.10 | +.02 | -.15 | -.07 | -.06 |
| 9. | -.59 | -.64 | -.68 | +.31 | +.05 | +.15 | +.37 | +.59 | +.35 | -.21 | +.08 | -.07 | +.03 | -.03 | -.05 |
| 10. | -.56 | -.60 | -.56 | +.52 | +.35 | +.30 | +.05 | +.16 | -.03 | +.09 | +.15 | 0 | -.02 | 0 | -.02 |
| 13. | -.74 | -.81 | -.66 | +.22 | +.09 | +.13 | +.02 | +.11 | -.08 | +.10 | +.18 | 0 | +.02 | +.01 | +.06 |
| 20. | +.63 | +.73 | +.62 | +.16 | +.05 | +.11 | +.08 | +.02 | +.14 | -.13 | -.28 | -.17 | +.05 | +.04 | -.01 |
| 22. | +.74 | +.86 | +.75 | +.21 | +.21 | +.31 | +.11 | +.02 | +.09 | -.14 | -.11 | -.04 | -.16 | -.05 | -.12 |
| 33. | +.64 | +.78 | +.61 | -.06 | -.15 | +.01 | +.28 | +.05 | +.26 | -.01 | +.04 | +.06 | +.12 | +.12 | +.02 |
| 39. | -.71 | -.84 | -.71 | 0 | -.03 | -.09 | -.04 | 0 | -.08 | +.02 | +.04 | -.04 | -.01 | -.01 | +.03 |
| 41. | +.48 | +.64 | +.46 | -.17 | -.31 | -.08 | +.45 | +.15 | +.44 | +.13 | +.07 | +.10 | +.06 | +.07 | -.04 |
| 11. | -.44 | -.60 | -.43 | -.71 | -.39 | -.80 | -.20 | -.23 | -.08 | -.21 | -.25 | 0 | +.23 | -.01 | +.01 |
| 12. | +.13 | +.11 | -.03 | +.86 | +.66 | +.78 | +.01 | +.15 | -.01 | +.03 | +.13 | -.02 | +.05 | +.11 | +.18 |
| 14. | +.32 | +.25 | +.38 | -.67 | -.52 | -.50 | +.15 | -.14 | +.16 | -.09 | -.18 | -.01 | +.14 | +.12 | +.03 |
| 15. | +.28 | +.33 | +.41 | -.62 | -.52 | -.49 | -.11 | -.32 | +.02 | +.27 | +.02 | +.19 | +.04 | +.04 | +.03 |
| 21. | +.06 | +.13 | +.01 | +.85 | +.77 | +.77 | -.21 | -.03 | -.20 | +.10 | +.02 | +.11 | 0 | +.03 | -.04 |
| 31. | -.21 | -.09 | -.20 | +.77 | +.51 | +.67 | +.14 | +.31 | 0 | +.08 | +.24 | +.15 | 0 | -.03 | -.05 |
| 35. | +.03 | +.12 | +.15 | -.35 | -.57 | -.63 | +.45 | +.20 | +.18 | -.23 | -.11 | -.05 | -.02 | -.03 | -.17 |
| 37. | +.16 | +.44 | +.09 | +.82 | +.63 | +.81 | +.04 | +.18 | +.10 | -.05 | -.07 | +.06 | +.09 | +.09 | 0 |
| 42. | -.08 | -.23 | -.24 | +.64 | +.62 | +.51 | -.20 | -.03 | -.07 | -.35 | -.12 | -.13 | +.40 | +.42 | +.48 |
| 5. | -.02 | -.21 | -.20 | +.09 | +.46 | +.19 | -.80 | -.62 | -.69 | -.05 | -.04 | -.05 | 0 | +.01 | +.13 |
| 8. | +.09 | -.03 | 0 | -.52 | -.21 | -.32 | -.57 | -.64 | -.65 | +.02 | -.06 | -.06 | -.03 | -.10 | 0 |
| 16. | -.08 | -.10 | +.13 | -.17 | -.20 | -.20 | +.46 | +.59 | +.64 | -.50 | -.32 | -.18 | -.06 | -.06 | -.14 |
| 18. | +.33 | +.39 | +.34 | -.06 | -.40 | -.11 | +.67 | +.51 | +.57 | +.03 | +.03 | +.05 | +.03 | +.05 | -.02 |
| 23. | +.04 | -.02 | +.13 | -.18 | -.54 | -.26 | +.60 | +.32 | +.62 | +.12 | +.05 | +.12 | +.37 | +.34 | +.18 |
| 34. | +.18 | +.05 | -.02 | -.14 | -.07 | -.11 | +.49 | +.45 | +.69 | -.38 | -.27 | -.16 | +.21 | +.27 | +.17 |
| 36. | -.01 | +.19 | +.07 | +.08 | +.01 | -.07 | +.53 | +.59 | +.68 | -.35 | -.20 | -.20 | +.10 | +.01 | -.16 |
| 38. | -.24 | +.01 | -.16 | +.35 | +.03 | +.37 | +.66 | +.74 | +.49 | -.01 | +.28 | -.05 | -.02 | -.04 | -.35 |
| 40. | +.07 | +.13 | +.04 | -.17 | +.20 | +.05 | -.71 | -.63 | -.71 | -.04 | +.11 | +.04 | 0 | +.04 | -.01 |
| 2. | -.31 | -.15 | -.31 | +.52 | +.20 | +.34 | +.09 | 0 | +.05 | +.54 | +.63 | +.55 | -.08 | +.04 | -.07 |
| 19. | -.21 | -.30 | -.18 | -.02 | -.12 | +.05 | -.06 | +.04 | -.21 | +.67 | +.68 | +.62 | -.23 | -.09 | -.26 |
| 24. | +.48 | +.36 | +.44 | -.09 | 0 | +.03 | +.04 | +.21 | +.15 | -.52 | -.54 | -.52 | -.14 | -.07 | -.10 |
| 27. | +.03 | +.16 | +.24 | 0 | -.08 | -.04 | -.15 | -.36 | -.14 | +.69 | +.54 | +.79 | +.04 | -.01 | -.07 |
| 29. | -.10 | -.10 | -.30 | +.03 | +.21 | +.12 | +.30 | +.08 | +.44 | -.16 | -.10 | +.05 | +.75 | +.80 | +.57 |
| 32. | +.38 | +.22 | +.23 | +.18 | -.01 | +.34 | +.02 | -.21 | +.05 | +.27 | +.19 | -.02 | +.31 | +.59 | +.59 |
| 3. | +.42 | +.58 | +.43 | -.13 | -.41 | -.15 | +.61 | +.34 | +.46 | +.09 | +.09 | +.16 | +.02 | +.02 | -.05 |
| 17. | +.08 | +.05 | -.05 | +.46 | +.68 | +.54 | -.45 | -.02 | -.29 | -.53 | -.57 | -.44 | -.06 | -.11 | -.11 |
| 25. | -.43 | -.58 | -.30 | -.34 | -.07 | -.32 | -.47 | -.43 | -.52 | -.03 | -.12 | -.10 | -.02 | -.09 | -.03 |
| 26. | -.31 | -.36 | -.23 | +.17 | +.21 | +.06 | -.40 | -.15 | -.54 | +.13 | +.10 | -.03 | -.40 | -.48 | -.25 |
| 30. | -.28 | -.24 | -.38 | +.44 | +.31 | +.40 | +.36 | +.47 | +.24 | -.13 | +.13 | -.09 | -.16 | -.14 | -.11 |
| % trace | 28 | 36 | 46 | 19 | 21 | 20 | 9 | 13 | 7 | 7 | 5 | 6 | 7 | 4 | 4 |

Tab

**Prediction of Year 1 Grade Point Average with a Battery of 72 Variables:**

**42 Original Peer, 13 Aptitude, 2 High School, and 15 Edwards Variables**

| Name of Variable | Univariate Correlation | Beta Weights | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Steps in analysis | | | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Adaptable (Peer) | +.12 | | | | | | | | | | +.12 |
| Conscientious (Peer) | +.32 | | | | | | | +.18 | +.17 | +.15 | +.13 |
| Quitting (Peer) | -.47 | -.47 | -.47 | -.47 | -.45 | -.44 | -.49 | -.41 | -.35 | -.34 | -.32 |
| Happy-go-lucky(Peer) | -.16 | | | | | | | | | -.10 | -.14 |
| Considerate (Peer) | +.05 | | | | | | -.12 | -.21 | -.20 | -.19 | -.24 |
| Inquisitive (Peer) | +.35 | | | | | | | | +.11 | +.13 | +.15 |
| High School Rank | +.19 | | | | +.13 | +.17 | +.17 | +.16 | +.17 | +.16 | +.15 |
| SAT Verbal | +.22 | | | | | | +.16 | +.15 | +.14 | +.12 | +.13 | +.14 |
| High School Certified Units | +.18 | | | +.19 | +.16 | +.16 | +.16 | +.16 | +.15 | +.15 | +.15 |
| Cooperative English (Level) | +.23 | | +.23 | +.26 | +.27 | +.20 | +.19 | +.20 | +.19 | +.20 | +.20 |
| Multiple regression value (R) at each step of the analysis: | | .47 | .53 | .56 | .57 | .59 | .60 | .61 | .62 | .63 | .64 |
| Contribution to $R^2$ made by the entering variable: | | .2225 | .0549 | .0364 | .0167 | .0202 | .0126 | .0134 | .0091 | .0089 | .0091 |

Total contribution to $R^2$: peer = .2756, aptitude = .0751, high school = .0531, Edwards = 0

Percent Contribution to $R^2$; peer = 68%, aptitude = 19%, high school = 13%, Edwards = 0%

Table Prediction of Year 1 Grade Point Average with a Battery of 35 Variables:

5 Composite Peer, 13 Aptitude, 2 High School, and 15 Edwards Variables

| Name of Variable | Univariate Correlation | Beta Weights | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Steps in analysis | | | | | | | | | |
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Agreeableness (Peer) | +.01 | | | | | -.14 | -.14 | -.13 | -.13 | -.09 | -.09 |
| Strength of Character (Peer) | +.43 | +.43 | +.44 | +.43 | +.42 | +.46 | +.45 | +.45 | +.46 | +.49 | +.48 |
| Emotionality (Peer) | +.01 | | | | | | | | | -.12 | -.11 |
| High School Rank | +.19 | | | +.19 | +.16 | +.17 | +.17 | +.17 | +.17 | +.15 | +.13 |
| High School Certified Units | +.18 | | | | +.16 | +.16 | +.14 | +.15 | +.16 | +.16 | +.16 |
| Cooperative English (Level) | +.23 | | +.24 | +.26 | +.27 | +.26 | +.22 | +.21 | +.20 | +.21 | +.22 |
| Step Writing | +.25 | | | | | +.11 | +.12 | +.12 | +.11 | +.11 | |
| Order | -.05 | | | | | | | -.09 | -.11 | -.14 | |
| Nurturance | -.07 | | | | | | -.09 | -.11 | -.12 | -.12 | |
| Endurance | +.11 | | | | | | | | | +.10 | |
| Multiple regression value (R) at each step of the analysis: | | .43 | .49 | .53 | .55 | .56 | .57 | .58 | .59 | .60 | .60 |
| Contribution to $R^2$ made by the entering variable: | | .1876 | .0561 | .0360 | .0229 | .0161 | .0109 | .0072 | .0079 | .0105 | .0084 |
| Total contribution to $R^2$: peer = .2142, aptitude = .0670, high school = .0589, Edwards = .0235 | | | | | | | | | | | |
| Percent contribution to $R^2$: peer = 59%, aptitude = 18%, high school = 16%, Edwards = 6% | | | | | | | | | | | |

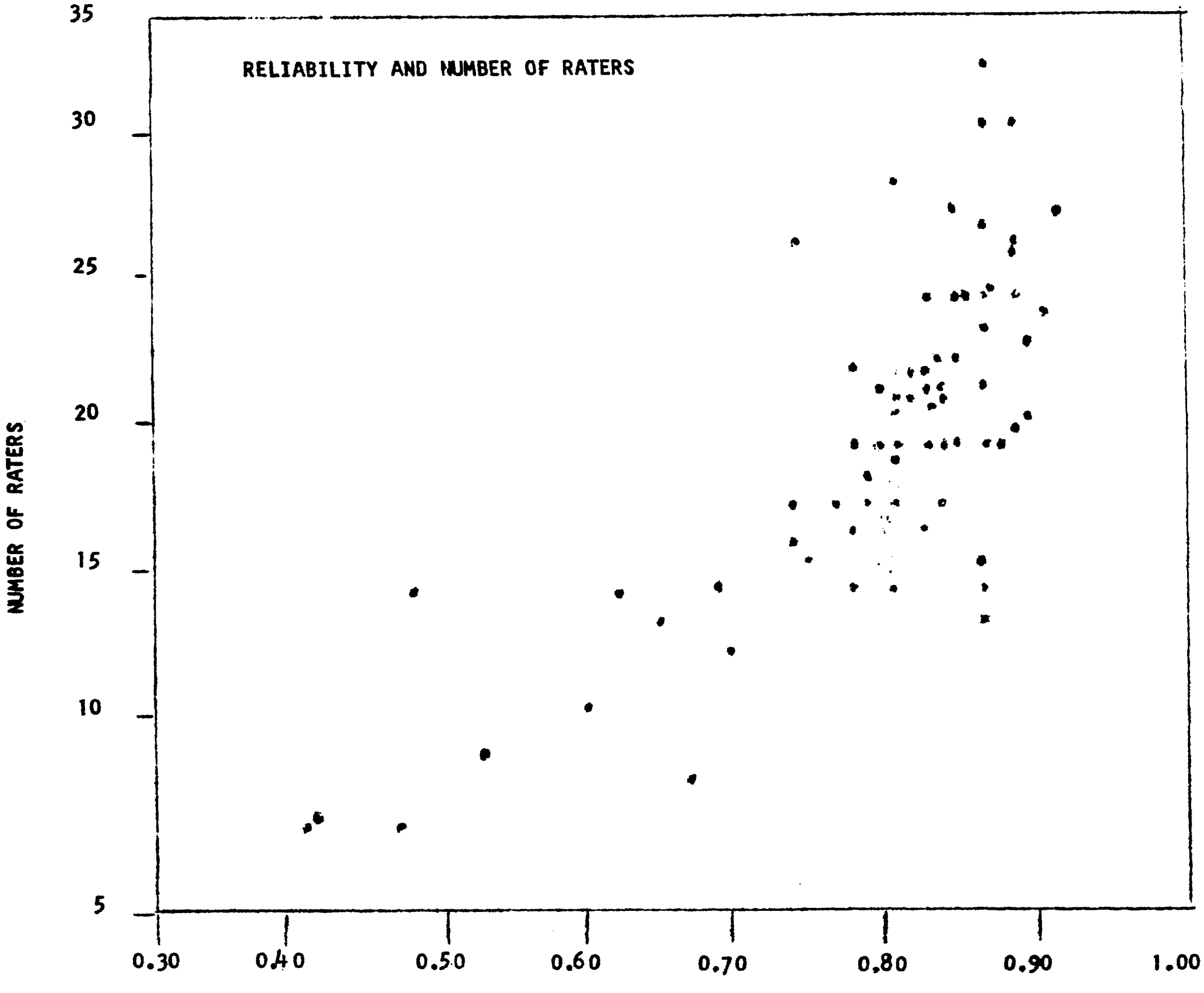
Figure 1

SAMPLE CARD USED FOR PEER
RATINGS OF PERSONALITY

| 1. ADAPTABLE: flexible; accepts changes of plan easily; satisfied with compromises; is not upset, surprised, baffled, or irritated if things are different from what he expected. | VS. | RIGID: insists that things be done the way he has always done them; does not adapt his habits and ways of thinking to those of the group; nonplussed if his routine is upset. |
|---|---|---|
| [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] | | [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] |
| [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] | | [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] |
| [21] [22] [23] [24] [25] [26] [27] [28] [29] [30] | | [21] [22] [23] [24] [25] [26] [27] [28] [29] [30] |
| [31] [32] [33] [34] [35] [36] [37] [38] [39] [40] | | [31] [32] [33] [34] [35] [36] [37] [38] [39] [40] |

Rating Instructions:

1. On left, mark identification numbers of the 5 most ADAPTABLE students in your group.
2. On right, mark identification numbers of the 5 most RIGID students in your group.


RELIABILITY AND NUMBER OF RATERS
NUMBER OF RATERS
35
30
25
20
15
10
5
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
MEDIAN RELIABILITY OF 42 PEER RATING TRAITS

## ACKNOWLEDGMENTS

This work was supported in part by grants #NU00104 and USPH-M-987 from the United States Public Health Service, in part by a grant from the Council for Tobacco Research-U.S.A., and in part by a grant from the Spaulding-Potter Charitable Trusts. Frederick Mosteller, Professor of Mathematical Statistics, Harvard University, and Arthur P. Dempster, Professor of Statistics, Harvard University, provided the author helpful guidance and this is gratefully acknowledged.